

Detecting and interpreting distortions in hierarchical organization of complex time series

Stanisław Drożdż^{1,2} and Paweł Oświęcimka^{1,*}

¹*Institute of Nuclear Physics, Polish Academy of Sciences, Kraków, Poland.*

²*Faculty of Physics, Mathematics and Computer Science,
Cracow University of Technology, Kraków, Poland.*

(Dated: March 10, 2015)

Hierarchical organization is a cornerstone of complexity and multifractality constitutes its central quantifying concept. For model uniform cascades the corresponding singularity spectra are symmetric while those extracted from empirical data are often asymmetric. Using the selected time series representing such diverse phenomena like price changes and inter-transaction times in the financial markets, sentence length variability in the narrative texts, Missouri River discharge and Sunspot Number variability as examples, we show that the resulting singularity spectra appear strongly asymmetric, more often left-sided but in some cases also right-sided. We present a unified view on the origin of such effects and indicate that they may be crucially informative for identifying composition of the time series. One particularly intriguing case of this later kind of asymmetry is detected in the daily reported Sunspot Number variability. This signals that either the commonly used famous Wolf formula distorts the real dynamics in expressing the largest Sunspot Numbers or, if not, that their dynamics is governed by a somewhat different mechanism.

PACS numbers: 05.10.-a, 05.45.Df, 05.45.Tp

Multi-scale approach [1–3] aims at bridging the wide range of time and length scales that are inherent in a number of essential processes in complex natural phenomena. Multifractality constitutes a leading concept towards quantifying the related characteristics [4]. By now it finds applications in essentially all areas of the scientific activity including physics [5, 6], biology [7–9], chemistry [10, 11], geophysics [12, 13], economics [14–21], hydrology [22], atmospheric physics [23], quantitative linguistics [24, 25], behavioral sciences [26], music [27, 28], and even ecological sciences [29].

At present the most efficient, numerically stable and precise [30] method to quantify multifractality is based on the Multifractal Detrended Fluctuation Analysis (MFDFA) [31]. Accordingly, for a discrete signal $x(i)_{i=1,\dots,N}$ one starts with the signal profile $X(j) = \sum_{i=1}^j (x(i) - \langle x \rangle)$, $j = 1, \dots, N$, where $\langle \dots \rangle$ denotes averaging over all i 's. Then one divides the $X(j)$ into M_s non-overlapping segments of length s ($s < N$) starting from both the beginning and the end of the signal (there are thus in total $2M_s$ segments). For each segment a local trend is estimated by fitting an l th order polynomial $P_\nu^{(l)}$, which is then subtracted from the signal profile. For the so-detrended signal a local variance $F^2(\nu, s)$ in each segment ν is calculated for the scale variable s . Finally, by averaging $F^2(\nu, s)$ over all segments ν one calculates the q th order fluctuation function:

$$F_q(s) = \left\{ \frac{1}{2M_s} \sum_{\nu=1}^{2M_s} [F^2(\nu, s)]^{q/2} \right\}^{1/q}, \quad (1)$$

and $q \in \mathbf{R}$. The optimal range is $q \in [-4, 4]$ [32].

The scaling behavior of $F_q(s) \sim s^{h(q)}$ indicates fractal structure with the singularity spectrum $f(\alpha)$ [1]

$$f(\alpha) = q[\alpha - h(q)] + 1 \quad (2)$$

for $\alpha = h(q) + qh'(q)$. If $h(q) = \text{const}$, the signal is monofractal. A nontrivial q -dependence of $h(q)$ indicates its more convoluted fractal organization termed multifractal. For model multifractal series $f(\alpha)$ typically assumes shape of an inverted symmetric parabola while the empirical ones often develop asymmetries. Here we explore this issue and document that asymmetries in $f(\alpha)$ may provide important information about the series organization.

Spectacular examples, largely novel, of unquestioned asymmetric multifractal spectra, generated using MFDFA algorithm, for signals representing diverse, mutually remote areas are shown in Fig. 1. They are seen to even display two distinct kinds of asymmetry. The DAX and DJIA stock market indices, the Forex market represented by the three exchange rate pairs (CHF/JPY, GBP/CHF and GBP/JPY), the series representing the sentence length variability in some selected literary texts (here *Rayuela* by Julio Cortázar and *2666* by Roberto Bolaño; most of the world famous literary works are essentially monofractal in the sentence length variability and just a dozen or so are convincingly multifractal and the two considered here at the same time belong to those few that develop the most asymmetric $f(\alpha)$) and the discharge of Missouri River [33] show pronounced left-sided asymmetry. A rarer, right-sided asymmetry is displayed by the other two examples, the stock market company (here ADS and BMW from DAX) inter-transaction times and the Sunspot Number variability [34]. In order to estimate a possible additional contribution to asymmetry

* pawel.oswiecimka@ifj.edu.pl

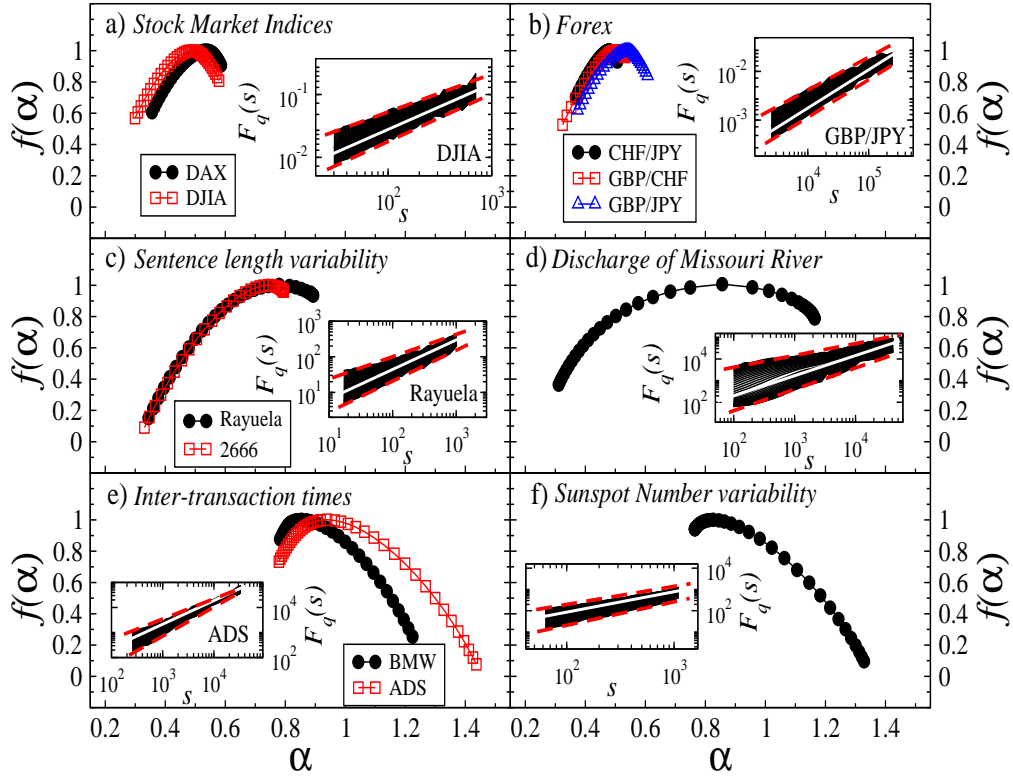


FIG. 1. (Color online) Singularity spectra $f(\alpha)$ for (a) daily returns of the two (DJIA and DAX) stock market indices over the period January 12, 1990 - October 12, 2013 (5881 data points), (b) minutely Forex returns on selected pairs (CHF/JPY, GBP/CHF, GBP/JPY) of currencies over the period 9:00 pm January 2, 2004 - 9:00 pm 20 March, 2008 (1703520 data points each), (c) the series representing length, expressed in terms of the number of words, of the consecutive sentences in the two narrative texts: *Rayuela* by Julio Cortázar (9848 sentences) and *2666* by Roberto Bolaño (21319 sentences), (d) hourly changes of the Missouri River discharge (in cubic feet per second) at Waverly, taken from the U.S. Geological Survey database [33] over the period October 1, 1987 - October 1, 2007 (213344 data points), (e) inter-transaction intervals over the period November 28, 1997 - December 31, 1999, for the two German stock market companies: ADS (167208 data points) and BMW (173664 data points) (f) daily Sunspot Number over the period January 1, 1900 - June 30, 2014 (41817 data points); Source: WDC-SILSO, Royal Observatory of Belgium, Brussels [39]. The insets show the corresponding $F_q(s)$ for $q \in [-4, 4]$ with the white strip indicating $q = 0$. The detrending polynomial used, as optimal, is of the second order. In case (f) the upper limit of the scaling parameter s is less than half of the 11 years solar activity cycle which makes this procedure unbiased by the periodic trend.

of the potential 'oscillating singularities' [35] we also applied the so-called wavelet leaders algorithm [36, 37] to all these data but detected no significant signals of such oscillations.

The left side of $f(\alpha)$ is determined by the positive q -values, which filter out larger events and the opposite applies to its right side. Hence, asymmetry in $f(\alpha)$ signals non-uniformity of the underlying cascade. The (a)-(d) cases in Fig. 1 are thus seen to be more multifractal in arrangement of the large events and far less such in the small ones. In some of those cases (CHF/JPY, GBP/CHF or 2666) the right-side of $f(\alpha)$ contracts so strongly that it indicates essentially a monofractal character of the corresponding small fluctuations. This suggests that the entire signal can be considered a mixture of the large scale homogeneous multiplicative cascade, giving thus rise to a pronounced left wing in $f(\alpha)$, and of the small scale noise-like background, shrinking the corresponding right wing in $f(\alpha)$. Such a noise-like

component in empirical data is not very unusual as it may for instance originate from the measurement uncertainty or from some coarse graining that effectively affects more just the small scale fluctuations. The reverse kind of asymmetry - right-sided - applies, however, to the cases (e) and (f), and here the situation is more intriguing and unusual since it indicates that pronounced multifractality operates on small scale fluctuations while the dynamics of large scale fluctuations is much poorer in this respect. That the stock market company inter-transaction intervals may be governed by such a dynamics is conceivable. Small intervals occur between transactions grouped within the same clusters of an enhanced volatility on this particular company and may thus be strongly, also nonlinearly correlated. The distances between such in time more separated clusters of activity, thus larger inter-transaction intervals connecting them, may be less correlated [32].

More difficult to interpret - thus even more interesting

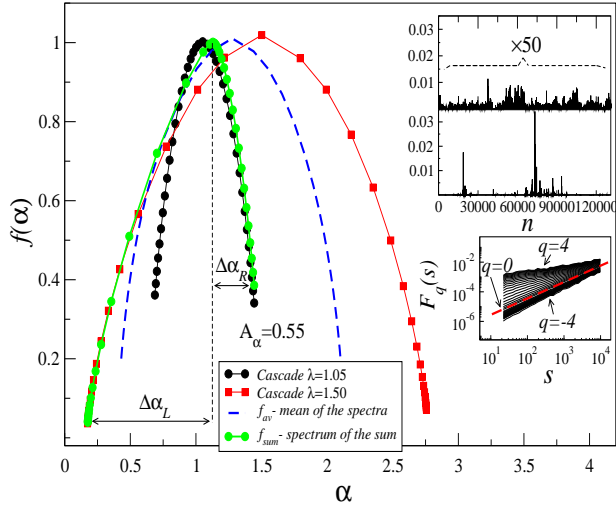


FIG. 2. (Colore online) Examples of $f(\alpha)$ for the two binomial cascades with $\lambda = 1.05$ and $\lambda = 1.5$, the average $f_{av}(\alpha)$ of such two $f(\alpha)$'s and $f_{sum}(\alpha)$ calculated for the sum of these two cascades. The upper two insets show explicit spreads of these two cascades. To better reflect proportions in the same scale the original values of $\lambda = 1.05$ cascade are multiplied by a factor of 50. The lowest inset displays $F_q(s)$ functions for the sum of both cascades.

- is the analogous result for the Sunspot Number variability. The smaller fluctuations (seen through $q < 0$) develop a broad right-sided $f(\alpha)$ which suggests their cascade-like hierarchical organization. At the same time the large ones, filtered out by $q > 0$, show at most a remnant of multifractality. They thus do not belong to the same hierarchy or, at best, strongly distort it. This observation may either demand revision of the famous Wolf formula ($R = k \cdot (N_s + 10N_g)$, N_s - number of spots, N_g - number of groups) [38] commonly used for expressing the Sun activity in terms of the number of spots or may indicate that the corresponding large and small activities are governed by a somewhat different cascading mechanisms.

The quality of scaling can be assessed from the insets to the corresponding panels of Fig. 1 which show the $F_q(s)$ dependencies on s . In (a) - (d) one in addition sees that the scaling exponents (slope of $F_q(s)$) visibly depend on q for $q > 0$ while very weakly for $q < 0$. The opposite is true for (e) and (f) cases. Naturally, this correlates with orientations of asymmetries seen in $f(\alpha)$.

The most likely origin and thus interpretation of the above effects can transparently be illustrated by using the model cascades. Here we use the binomial cascade with multipliers whose logarithms are drawn from the Gaussian $N(\lambda, 1)$ distribution. Examples of $f(\alpha)$ for two such cascades with $\lambda = 1.05$ and $\lambda = 1.5$, the average $f_{av}(\alpha)$ of such two $f(\alpha)$'s and $f_{sum}(\alpha)$ calculated for the sum of these two cascades are shown in the main panel of Fig. 2. While $f_{av}(\alpha)$ is symmetric, the $f_{sum}(\alpha)$ is much narrower and already strongly asymmetric. To quantify these characteristics we introduce the asymmetry param-

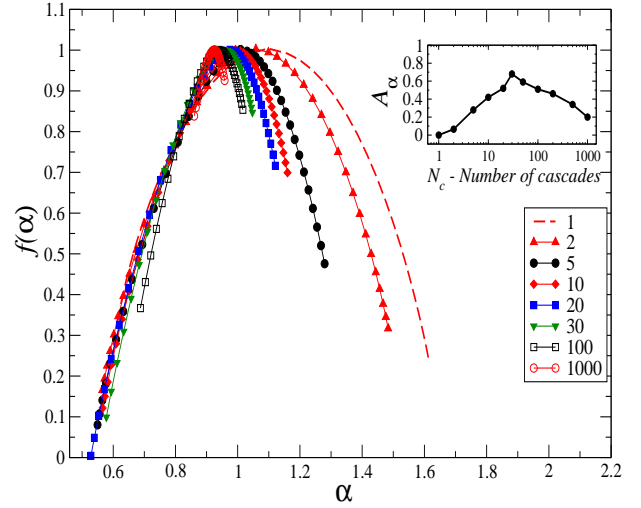


FIG. 3. (Color online) Singularity spectra of $f(\alpha)$ for an increasing number of the superimposed binomial cascades generated independently but all with the same $\lambda = 1.1$. Inset illustrates how the corresponding asymmetry parameter A_α evolves.

eter

$$A_\alpha = (\Delta\alpha_L - \Delta\alpha_R)/(\Delta\alpha_L + \Delta\alpha_R) \quad (3)$$

where $\Delta\alpha_L = \alpha_0 - \alpha_{min}$ and $\Delta\alpha_R = \alpha_{max} - \alpha_0$ and α_{min} , α_{max} , α_0 denote the beginning and the end of $f(\alpha)$ support, and the α value at maximum of $f(\alpha)$ (which corresponds to $q = 0$), respectively. A simple superposition of only two such stochastic binomial cascades with different parameters produces strongly left-sided asymmetric ($A_\alpha = 0.55$) multifractal spectrum $f(\alpha)$, resembling the corresponding empirical ones from Fig. 1. It is instructive to see that the left side of $f_{sum}(\alpha)$ coincides with $f(\alpha)$ for the $\lambda = 1.5$ cascade while its right side with the one for $\lambda = 1.05$ cascade. The origin of this result becomes clear from the relative proportions of the size of fluctuations in these two individual cascades as seen in the upper two panels in Fig. 2. The $\lambda = 1.05$ fluctuations are on average about two orders of magnitude smaller than the $\lambda = 1.5$ ones. Thus in their sum the small fluctuations are dominated by $\lambda = 1.05$ and filtered out by $q < 0$ which corresponds to the right side in f_{sum} and the opposite applies to large fluctuations. An equivalent way to see this effect is through $F_q(s)$ defined by Eq. 1 and shown in the lowest panel of Fig. 2. Indeed, all $F_q(s)$ scale with s but spread much more for $q > 0$, thus width on the left side of $f_{sum}(\alpha)$ is larger.

Similar effect of left-sided asymmetry one even obtains by superimposing cascades generated with the same λ . The result for an increasing number N of independently drawn and then superimposed cascades with $\lambda = 1.1$ is displayed in Fig. 3. Here however more components are needed to reach the same degree of asymmetry as in the Fig. 2 case. The left-sided asymmetry originates here from the fact that small fluctuations are more abundant

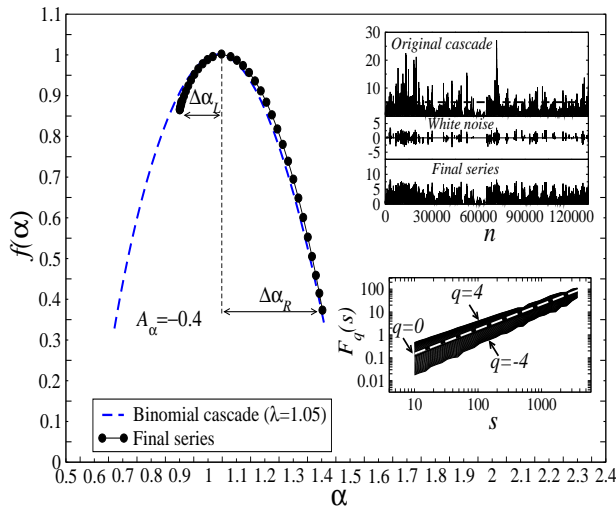


FIG. 4. (Color online) Singularity spectrum $f(\alpha)$ (i) for $\lambda = 1.05$ binomial cascade and (ii) for the one obtained from it by setting a threshold of 4.7, then from all events that exceed this threshold value subtracting the over-threshold value (which affects 1% of all events) and finally by adding consecutively to those affected one number drawn from a normal distribution, one independently to each. Upper inset, from top to bottom, shows the original binomial cascade with the truncation threshold (dashed line), the white noise added and the resulting final series. Lower inset displays the q -dependence of the fluctuation functions $F_q(s)$ for the so-obtained final series.

in such individual cascades therefore summing up an increasing number of them one sooner approaches the white noise limit, i.e., destroys the original hierarchical organization on the level of small fluctuations than on the level of the large ones. The right side of the resulting $f(\alpha)$ contracts sooner to a monofractal with an increasing N than does the corresponding left side. This can be seen from Fig. 3, especially from the inset which shows the A_α dependence on N . After the initial sharp increase it reaches maximum at around $N = 30$ with $A_\alpha \approx 0.7$ and then slowly starts decreasing. For $N = 1000$ the $f(\alpha)$ already is much closer to a monofractal with $A_\alpha \approx 0.2$.

While the left-sided multifractal asymmetry easily emerges when superimposing cascades, the right-sided one is more peculiar to model. From an algorithmic perspective it implies just a reverse, i.e., a more uniform hierarchical organization on the level of smaller fluctuations and more noise-like behavior of the large fluctuations. Such an interpretation indicates construction of a series obeying condition of this kind. An example is displayed in Fig. 4. This is the binomial cascade generated from the same $\lambda = 1.05$ model as before, but now a top section of the largest events is randomized as follows. 1% of the largest events is selected which in this case corresponds to those that exceed a threshold of $T = 4.7\sigma$. Then, their values are replaced by the sum of T and of a number drawn independently for each event from $N(0, 1)$.

The net result of such a construction preserves organization of the original small fluctuations and randomizes the largest ones. The so obtained series is displayed in the upper inset of Fig. 4 (*final series*), together with the two components that are used to form it. The corresponding $F_q(s)$ functions (lower panel in Fig. 4) become now more spread for $q < 0$, thus the $f(\alpha)$ spectrum becomes right-sided ($A_\alpha = -0.4$), which confirms the demanded reversed organization. Of course, by playing with T or with noise added one may modulate the degree of this right-sided asymmetry.

In conclusion the effects of asymmetry in singularity spectra $f(\alpha)$ of the time series representing complex systems may not just be numerical artifacts as they seem to be treated in the literature on the subject, but instead may contain genuine information about the composition of such series. They then indicate at least two different regimes in the multifractal scaling. The left-sided asymmetry corresponds to a pronounced multifractality on the level of larger fluctuations and its suppression towards monofractality when going to small fluctuations. These small ones thus resemble a noisy background accompanying the entire signal. Globally, such structures can easily be modeled as a superposition of independent, multifractally symmetric cascades. Even in relation to reality such a modeling seems appropriate since participation of several somewhat independent multiplicative factors on a given phenomenon seems indeed possible. Out of those studied here the stock market index is an explicit sum of prices of the individual companies entering the basket. The span of $f(\alpha)$ of such sums appears also visibly narrower. This thus calls for caution when interpreting the width of $f(\alpha)$ as a measure of the 'degree of complexity'. In case of asymmetric $f(\alpha)$ application of an appropriate multifractal cross-correlation analysis [40–42] in order to disentangle the original signal into uniform cascade components may in this connection be recommended. In the right-sided asymmetry the situation is reversed. Here, these are the smaller fluctuations that develop a pronounced multifractal hierarchy while the largest ones are suppressed. In relation to natural phenomena this seems to be a much rarer category of distortion. Especially intriguing is the Sunspot Number variability case belonging to this category. The arguments presented here indicate that possibly the famous Wolf formula [38] commonly used to digitally express this aspect of the Sun activity does not represent a faithful mapping for the large events and needs to be refined such that the character of cascading from the smaller events be extended to the large ones. Of course, at the present stage another possibility that the large events are governed by a somewhat different dynamics cannot be excluded either. In any case, resolving this issue emerges as a great scientific challenge.

We thank Drs Frederic Clette and Laure Lefevre of the Royal Observatory of Belgium, Brussels, for very helpful exchanges on the issue of the Sunspot Number variability.

-
- [1] T.C. Halsey, M.H. Jensen, L.P. Kadanoff, I. Procaccia, and B.I. Shraiman, *Phys. Rev. A* **33**, 1141 (1986).
 - [2] B. B. Mandelbrot, *Pure Appl. Geophys.* **131**, 5 (1989).
 - [3] J.-F. Muzy, E. Bacry, and A. Arneodo, *Int. J. Bifur. Chaos* **4**, 245 (1994).
 - [4] J. Kwapien, and S. Drożdż, *Phys. Rep.* **515**, 115-226 (2012).
 - [5] J.F. Muzy, E. Bacry, R. Baile, and P. Poggi, *EPL* **82**, 60007 (2008).
 - [6] A.R. Subramaniam, I.A. Gruzberg, and A.W.W. Ludwig, *Phys. Rev. B* **78**, 245105 (2008).
 - [7] P.Ch. Ivanov, L. A. N. Amaral, A. L. Goldberger, S. Havlin, M. G. Rosenblum, Z. R. Struzik, and H. E. Stanley, *Nature* **399**, 461-465 (1999).
 - [8] D. Makowiec, A. Dudkowska, R. Gałaska, and A. Rynkiewicz, *Physica A* **388** 3486-3502 (2009).
 - [9] A. Rosas, E. Nogueira Jr., and J.F. Fontanari, *Phys. Rev. E* **66**, 061906 (2002).
 - [10] H.E. Stanley and P. Meakin, *Nature* **335**, 405-409 (1988).
 - [11] V.V. Udovichenko and P.E. Strizhak, *Theor. Exp. Chem.* **38**, 259 (2002).
 - [12] A. Witt and B.D. Malamud, *Surv. Geophys.* **34**, 541 (2013).
 - [13] L. Telesca, V. Lapenna, and M. Macchiato, *New J. Phys.* **7**, 214 (2005).
 - [14] M. Ausloos and K. Ivanova, *Comp. Phys. Comm.* **147**, 582 (2002).
 - [15] L. Calvet and A. Fisher, *Rev. Econ. Stat.* **84**, 381 (2002).
 - [16] A. Turiel and C.J. Perez-Vicente, *Physica A* **355**, 475 (2005).
 - [17] S. Drożdż, J. Kwapien, P. Oświęcimka, and R. Rak, *New J. Phys.* **12** 105003 (2010).
 - [18] P. Oświęcimka, J. Kwapien, S. Drożdż, A.Z. Górski, and R. Rak, *Acta Phys. Pol. A* **114**, 547 (2008).
 - [19] W.-X Zhou, *EPL* **88**, 28004 (2009).
 - [20] M.I. Bogachev and A. Bunde, *Phys. Rev. E* **80**, 026131 (2009).
 - [21] Z.-Y. Su, Y.-T. Wang and H.-Y. Huang, *J. Korean Phys. Soc.* **54**, 1395 (2009).
 - [22] E. Koscielny-Bunde, J.W. Kantelhardt, P. Braund, A. Bunde and S. Havlin, *J. Hydrology* **322**, 120 (2006).
 - [23] J.W. Kantelhardt, E. Koscielny-Bunde, D. Rybski, P. Braun, A. Bunde and S. Havlin, *J. Geophys. Res. (Atmosph.)* **111**, D01106 (2006).
 - [24] M. Ausloos, *Phys. Rev. E* **86**, 031108 (2012).
 - [25] I. Grabska-Gradzińska, A. Kulig, J. Kwapien, P. Oświęcimka and S. Drożdż, *AWERProcedia Information Technology & Computer Science* **03**, 1700 (2013).
 - [26] E.A.F. Ihlen and B. Vereijken, *Human Mov. Sci.* **32** (2013) 633651.
 - [27] G.R. Jafari, P. Pedram, and L. Hedayatifar, *J. Stat. Mech.*, P04012 (2007).
 - [28] P. Oświęcimka, J. Kwapien, I. Celińska, S. Drożdż, and R. Rak, *arXiv:1106.2902* (2011).
 - [29] D.G. Kelty-Stephen, K. Palatinus, E. Saltzman and J.A. Dixon, *Ecological Psychology* **25** (2013) 1-62.
 - [30] P. Oświęcimka, J. Kwapien, and S. Drożdż, *Phys. Rev. E* **74**, 016103 (2006).
 - [31] J.W. Kantelhardt, S. A. Zschiegner, E. Koscielny-Bunde, S. Havlin, A. Bunde, and H.E. Stanley, *Physica A* **316**, 87 (2002).
 - [32] S. Drożdż, J. Kwapien, P. Oświęcimka, and R. Rak, *EPL* **88**, 60003, (2009).
 - [33] <http://ida.water.usgs.gov>
 - [34] R.J. Bray and R.E. Loughhead, *Sunspots* (Dover Publications, New York, 1979).
 - [35] S. Seuret, *Math. Nachr.* **279**, 1195 (2006).
 - [36] B. Lashermes, S.G. Roux, P. Abry, and S. Jaffard, *Eur. Phys. J. B* **61**, 201 (2008).
 - [37] E. Serrano, and A. Figliola, *Physica A* **388**, 2793 (2009).
 - [38] J.R. Wolf, *Mittheil. Sonnen. Nr.* 9 (1859)
 - [39] SILSO, World Data Center - Sunspot Number and Long-term Solar Observations, Royal Observatory of Belgium, on-line Sunspot Number catalogue: <http://www.sidc.be/SILSO/>, year(s)-of-data: January 1, 1900 - June 30, 2014
 - [40] B. Podobnik and H.E. Stanley, *Phys. Rev. Lett.* **100**, 084102 (2008).
 - [41] W.-X Zhou, *Phys. Rev. E* **77**, 066211 (2008).
 - [42] P. Oświęcimka, S. Drożdż, M. Forczek, S. Jadach and J. Kwapien, *Phys. Rev. E* **89**, 023305 (2014)